Combinatorics, Probability and Computing (2025), **34**, pp. 326–337
doi:10.1017/S0963548324000257

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Essential covers of the hypercube require many hyperplanes

Lisa Sauermann[1,*] and Zixuan Xu[2]

[1]University of Bonn, Bonn, Germany and [2]Massachusetts Institute of Technology, Cambridge, USA
Corresponding author: Zixuan Xu; Email: zixuanxu@mit.edu

(Received 30 November 2023; revised 11 June 2024; accepted 24 July 2024; first published online 16 December 2024)

Abstract

We prove a new lower bound for the almost 20-year-old problem of determining the smallest possible size of an essential cover of the n-dimensional hypercube $\{\pm 1\}^n$, that is, the smallest possible size of a collection of hyperplanes that forms a minimal cover of $\{\pm 1\}^n$ and such that, furthermore, every variable appears with a non-zero coefficient in at least one of the hyperplane equations. We show that such an essential cover must consist of at least $10^{-2} \cdot n^{2/3}/(\log n)^{2/3}$ hyperplanes, improving previous lower bounds of Linial–Radhakrishnan, of Yehuda–Yehudayoff, and of Araujo–Balogh–Mattos.

Keywords: Essential covers; hypercube

2020 MSC Codes: Primary: 05D40

1. Introduction

There is a long line of research, spanning over three decades, on problems about covering the vertices of the n-dimensional hypercube $\{\pm 1\}^n$ by hyperplanes (see, e.g. [2, 3, 6, 9, 10, 12]). A very simple question is how many hyperplanes are needed in order to cover all vertices in $\{\pm 1\}^n$, that is, such that every vertex is contained in at least one of these hyperplanes. This question has an equally simple answer: two hyperplanes are enough, for example, one can take the two hyperplanes given by the equations $x_1 = 1$ and $x_1 = -1$. However, this hyperplane cover is not truly n-dimensional, in the sense that the variables $x_2, \dots, x_n$ do not appear in any of the hyperplane equations. It is therefore natural to demand in addition that each variable has a non-zero coefficient in the equation for at least one of the hyperplanes.

However, this still does not lead to a very interesting problem. In addition to the hyperplanes with equations $x_1 = 1$ and $x_1 = -1$ that already cover all vertices, one could add an additional hyperplane with an equation like $x_1 + x_2 + \cdots + x_n = 0$ containing all variables with non-zero coefficients. But in some sense, this additional hyperplane is not truly part of the hyperplane cover, as it is not needed for covering all vertices (the other two hyperplanes together already cover all vertices in $\{\pm 1\}^n$).

This leads to the following notion of an *essential cover* of the hypercube $\{\pm 1\}^n$, introduced by Linial and Radhakrishnan [10] in 2005 and further studied in [3, 11, 12]. It describes a minimal set of hyperplanes covering all vertices in $\{\pm 1\}^n$, which in addition has the property that each variable appears in the equation of at least one of the hyperplanes. Formally, an essential cover of the hypercube $\{\pm 1\}^n$ is defined as follows.

*Research supported in part by NSF Award DMS-2100157 and a Sloan Research Fellowship.

© The Author(s), 2024. Published by Cambridge University Press.

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

Definition 1.1 (Essential cover). *A collection of hyperplanes $h_1, \dots, h_k$ in $\mathbb{R}^n$ is called an essential cover of the n-dimensional hypercube $\{\pm 1\}^n$ if the following conditions are satisfied:*

- *(E1) For every vertex $x \in \{\pm 1\}^n$, there is some $i \in [k]$ such that $x \in h_i$.*
- *(E2) For every $j \in [n]$, there is some $i \in [k]$ such that, writing $h_i = \{x \in \mathbb{R}^n \mid \langle v_i, x\rangle = \mu_i\}$, we have $v_{ij} \neq 0$.*
- *(E3) For every $i \in [k]$, there is a vertex $x \in \{\pm 1\}^n$ that is covered only by h_i and not by any of the hyperplanes $h_1, \dots, h_{i-1}, h_{i+1}, \dots, h_n$.*

Condition (E1) means that every vertex in $\{\pm 1\}^n$ is covered by some hyperplane, and condition (E3) means that $h_1, \dots, h_k$ is a minimal collection of hyperplanes with respect to the property of covering every vertex. Finally, condition (E2) can be restated as saying that each of the variables $x_1, \dots, x_n$ must appear with a non-zero coefficient in at least one of the hyperplane equations $\langle v_i, x\rangle = \mu_i$ for $i = 1, \dots, n$ (where $v_i = (v_{i1}, \dots, v_{in}) \in \mathbb{R}^n$ and $\mu_i \in \mathbb{R}$, so $\langle v_i, x\rangle = \mu_i$ is a shorter form of writing the linear equation $v_{i1}x_1 + \cdots + v_{in}x_n = \mu_i$). Note that this condition is not affected by rescaling the equations $\langle v_i, x\rangle = \mu_i$, so it does not depend on the parametrisation chosen for each of the hyperplanes.

It is now a very natural question to ask how large an essential cover of $\{\pm 1\}^n$ needs to be. Linial and Radhakrishnan [10] showed in 2005 that any essential cover must contain at least $\Omega(\sqrt{n})$ hyperplanes. On the other hand, they also gave a construction of an essential cover consisting of $\lceil n/2\rceil + 1$ hyperplanes, which still remains the best known upper bound. Yehuda and Yehudayoff [12] improved the lower bound to $\Omega(n^{0.52})$, and more recently Araujo, Balogh, and Mattos [3] further improved the lower bound to $\Omega(n^{5/9}/(\log n)^{4/9})$ by refining the methods in [12].

In this paper, we further improve the lower bound for the size of an essential cover of the hypercube.

Theorem 1.2. *For $n \geq 2$, any essential cover of the n-dimensional hypercube $\{\pm 1\}^n$ must consist of at least $10^{-2} \cdot n^{2/3}/(\log n)^{2/3}$ hyperplanes.*

The absolute constant 10^{-2} is not optimised in our proof.

We remark that our lower bound for the size of essential covers also implies new lower bounds for some problems in proof complexity, see the discussion in [12].

Linial and Radhakrishnan [10], obtained their lower bound of $\Omega(\sqrt{n})$ for the size of essential covers of $\{\pm 1\}^n$ by showing the following. For any essential cover of $\{\pm 1\}^n$ consisting of k hyperplanes, each hyperplane equation contains at most $2k$ variables with non-zero coefficients. We also use this fact in our proof of Theorem 1.2.

We furthermore observed that this fact leads to an essentially tight lower bound for another well-known problem concerning hyperplane covers of $\{\pm 1\}^n$, which can be stated as follows. A *skew cover* of the n-dimensional hypercube $\{\pm 1\}^n$ is a collection of hyperplanes $h_1, \dots, h_k$ covering all vertices in $\{\pm 1\}^n$ such that each of the hyperplane equations contains all of the variables $x_1, \dots, x_n$ with non-zero coefficients (i.e. for $i = 1, \dots, n$, when writing $h_i = \{x \in \mathbb{R}^n \mid \langle v_i, x\rangle = \mu_i\}$, the vector v_i has full support). The problem is again to determine the smallest possible size of a skew cover of $\{\pm 1\}^n$. The best lower bound for this problem that appeared explicitly in the literature was $\Omega(n^{0.51})$ due to Yehuda–Yehudayoff [13], but a bound of $\Omega(n^{2/3}/(\log n)^{4/3})$ follows from a recent result of Klein [9] via the same deduction as in [13]. From the above-mentioned fact about essential covers proved by Linial and Radhakrishnan [10], one can deduce the following stronger lower bound for the size of skew covers of $\{\pm 1\}^n$.

Observation 1.3. *For $n > 0$, any skew cover of the n-dimensional hypercube $\{\pm 1\}^n$ must consist of at least $n/2$ hyperplanes.*

We include a proof of this observation in Section 2, since this does not seem to have been previously observed in the literature (after the preprint version of this paper appeared, this observation

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

was also independently made in [8]). The bound in Observation 1.3 is tight up to constant factors, as it is not hard to construct a skew cover of $\{\pm 1\}^n$ consisting of $n+1$ hyperplanes (by taking the hyperplanes described by the equations $x_1 + \cdots + x_n = n - 2t$ for $t = 0, \ldots, n$).

Notation. For a positive integer n, we use $[n]$ to denote the set $\{1, \ldots, n\}$. Given sets A, B, C, we write $C = A \sqcup B$ to denote a partition of the set C into two disjoint subsets A and B. For any real $x > 0$, we write $\log x$ for the natural logarithm of x in base e. For two sets A, B, we use B^A to denote the set of maps $A \to B$. For $x \in B^A$ and a subset $A' \subseteq A$, we denote the restriction of x to A' by $x|_{A'}$, as usual. Given $x \in B^A$, one may think of x as a vector of length $|A|$ indexed by the elements in A with each coordinate taking values in B, and for $a \in A$, we write x_a for the value $x(a) \in B$. For convenience, given a positive integer n, we use B^n to denote $B^{[n]}$ (of course, this is consistent with the usual definition $B^n = B \times B \times \cdots \times B$).

For a vector $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, we use $\mathrm{supp}(v)$ to denote the support of v, which is the set of indices $i \in [n]$ where $v_i \neq 0$. We recall that the ℓ_1-norm of v is defined as $||v||_1 = \sum_{i=1}^n |v_i|$, the ℓ_2-norm of v is defined as $||v||_2 = \left(\sum_{i=1}^n v_i^2\right)^{1/2}$, and the ℓ_∞-norm of v is defined as $||v||_\infty = \max_{i \in [n]} |v_i|$. Given two vectors $u, v \in \mathbb{R}^n$, we write $\langle u, v \rangle = \sum_{i=1}^n u_i v_i$ for the standard inner product on $\mathbb{R}^n$. For an $n \times m$ matrix $V \in \mathbb{R}^{n \times m}$ and subsets $A \subseteq [n]$ and $B \subseteq [m]$, we use $V[A \times B]$ to denote the $|A| \times |B|$ submatrix of V consisting of the rows in A and the columns in B.

2. Preliminaries

In this section, we discuss useful definitions and lemmas for our argument. We start with the following two definitions that characterise vectors with many different magnitudes.

Definition 2.1 (Magnitude). *For any $x \in \mathbb{R} \setminus \{0\}$, we say that x has magnitude $j \in \mathbb{Z}$ if $10^j \leq |x| < 10^{j+1}$.*

Definition 2.2. *For a positive integer $S > 0$, we say that a vector v has at least S magnitudes if there exist S non-zero entries in v with distinct magnitudes.*

We remark that the notion of having many magnitudes is a simpler version of the notion of having "many scales" introduced in [13] and later used in [3, 12].

The motivation behind the above definitions is the following lemma, which states that given a vector $v \in \mathbb{R}^n$ with at least S magnitudes, for a random vector $w \in \{\pm 1\}^n$, the probability of the event $\langle v, w \rangle = \alpha$ for any given value $\alpha \in \mathbb{R}$ is exponentially small in S. For our purposes, it is convenient to state the lemma for a biased random vector $w \in \{\pm 1\}^n$. We remark that a similar lemma also appears in [9].

Lemma 2.3. *Let $v \in \mathbb{R}^n$ be a vector with at least S magnitudes. Let $p \in [1/3, 2/3]^n$ and consider a random vector $w \in \{\pm 1\}^n$ with independent random entries $w_1, \ldots, w_n \in \{\pm 1\}$, whose distributions are given by*

$$w_i = \begin{cases} 1 & \text{with probability } p_i \\ -1 & \text{with probability } 1 - p_i \end{cases} \qquad \text{for all } i \in [n].$$

Then for any $\alpha \in \mathbb{R}$, we have

$$\mathbb{P}[\langle v, w \rangle = \alpha] \leq \left(\frac{2}{3}\right)^{\lceil S/2 \rceil}.$$

Proof. Let $r = \lceil S/2 \rceil$. Among the at least S different magnitudes occurring for the entries of the vector v, there are at least r even numbers or at least r odd numbers. So, upon relabelling the indices, we may assume without loss of generality that the first r entries $v_1, \ldots, v_r$ of

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

$v=(v_1,\dots,v_n)\in\mathbb{R}^n$ have distinct magnitudes with the same parity and are ordered in decreasing order of magnitudes. Now, for each $i\in[r-1]$, if m is the magnitude of v_i, then v_{i+1} has magnitude at most $m-2$. This means that $|v_i|\geq 10^m$ and $|v_{i+1}|<10^{m-1}$, and therefore $|v_i|\geq 10|v_{i+1}|$ for every $i\in[r-1]$.

This implies that for any $\alpha\in\mathbb{R}$, there is at most one assignment for $w_1,\dots,w_r\in\{\pm 1\}$ satisfying $\sum_{i=1}^r v_i w_i=\alpha$. Indeed, if there were two distinct such assignments $(w_1,\dots,w_r),(w'_1,\dots,w'_r)\in\{\pm1\}^r$, then we would have $\sum_{i=1}^r v_i(w_i-w'_i)=\alpha-\alpha=0$. So for the minimum index $j\in\{1,\dots,r\}$ with $w_j\neq w'_j$, we would have

$$2|v_j|=\left|v_j(w_j-w'_j)\right|=\left|\sum_{i=j+1}^{r} v_i(w_i-w'_i)\right|\leq\sum_{i=j+1}^{r}2|v_i|\leq\sum_{i=j+1}^{r}2|v_j|/10^{j-i}<\frac{2}{9}\cdot|v_j|,$$

yielding a contradiction. So for any $\alpha\in\mathbb{R}$, there can be at most one such assignment $(w_1,\dots,w_r)\in\{\pm1\}^r$.

Let us now condition on arbitrary outcomes of $w_{r+1},\dots,w_n\in\{\pm1\}$. Conditional on such fixed outcomes, it suffices to show that, for any $\alpha\in\mathbb{R}$, we have $\sum_{i=1}^r v_i w_i=\alpha-\sum_{j=r+1}^n v_j w_j$ with probability at most $\left(\frac{2}{3}\right)^r$. Recall that there is at most one assignment for $w_1,\dots,w_r\in\{\pm1\}$ satisfying this equation. Since $w_1,\dots,w_r$ are independent random variables with

$$\mathbb{P}[w_i=1]\leq\frac{2}{3}\quad\text{and}\quad\mathbb{P}[w_i=-1]\leq\frac{2}{3}$$

for every $i\in[r]$, we obtain

$$\mathbb{P}\left[\sum_{i=1}^{r}v_i w_i=\alpha-\sum_{j=r+1}^{n}v_j w_j\,\middle|\,w_{r+1},\dots,w_n\right]\leq\left(\frac{2}{3}\right)^r=\left(\frac{2}{3}\right)^{\lceil S/2\rceil}.$$

This implies our desired statement. □

The following lemma is due to Linial and Radhakrishnan [10] and gives an upper bound for the number of non-zero coefficients for the hyperplane equations in an essential cover. It can be proved via the Combinatorial Nullstellensatz [1].

Lemma 2.4 ([10]). *Let $h_1,\dots,h_k$ be hyperplanes in $\mathbb{R}^n$ forming an essential cover of the hypercube $\{\pm1\}^n$. For each $i\in[k]$, let $h_i=\{x\in\mathbb{R}^n\mid\langle v_i,x\rangle=\mu_i\}$. Then we have $|\operatorname{supp}(v_i)|\leq 2k$ for all $i\in[k]$.*

We observed that Lemma 2.4 also gives a lower bound for the number of hyperplanes in any skew cover of the hypercube $\{\pm1\}^n$, as stated in Observation 1.3. Recall that the hyperplanes $h_1,\dots,h_k$ form a *skew cover* of the hypercube $\{\pm1\}^n$ if every vertex in $\{\pm1\}^n$ is covered by at least one of the hyperplanes $h_1,\dots,h_k$ and for each $i=1,\dots,k$, when writing $h_i=\{x\in\mathbb{R}^n\mid\langle v_i,x\rangle=\mu_i\}$, all coordinates of vector $v_i\in\mathbb{R}^n$ are non-zero.

Proof of Observation 1.3. Suppose $h_1,\dots,h_k$ are hyperplanes forming a skew cover of $\{\pm1\}^n$. For $i=1,\dots,k$, let us write $h_i=\{x\in\mathbb{R}^n\mid\langle v_i,x\rangle=\mu_i\}$. Then for every $i\in[k]$, all coordinates of the vector v_i are non-zero.

Let $I\subseteq[k]$ be a minimal subset with respect to the property that the hyperplanes h_i for $i\in I$ cover all vertices in $\{\pm1\}^n$ (and note that $I\neq\emptyset$). We claim that these hyperplanes h_i for $i\in I$ form an essential cover of $\{\pm1\}^n$. Indeed, conditions (E1) and (E3) in Definition 1.1 are satisfied by the choice of I. Condition (E2) is satisfied because for every index $i\in I$, all coordinates of v_i are non-zero (so for any $j\in[n]$, one can take an arbitrary index $i\in I$ in condition (E2)).

Now, taking an arbitrary index $i\in I$, Lemma 2.4 implies that $n=|\operatorname{supp}(v_i)|\leq 2|I|\leq 2k$, so we must have $k\geq n/2$. □

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

Our proof of Theorem 1.2 also relies on the following lemma, which Ball [4] extracted from Bang's solution of Tarski's plank problem [5] (see, e.g. [3] for a short self-contained proof of this lemma).

Lemma 2.5 (Bang's lemma [4, 5]). *Let M be a symmetric $\ell \times \ell$ matrix such that $M_{ii} = 1$ for every $i \in [\ell]$. Then for any vector $\mu = (\mu_1, \dots, \mu_\ell) \in \mathbb{R}^\ell$ and any real number $\theta \geq 0$, there exists $\varepsilon \in \{\pm 1\}^\ell$ such that*

$$\big|(M(\theta\varepsilon))_i - \mu_i\big| \geq \theta \quad \text{for all } i \in [\ell].$$

In our proof of Theorem 1.2, we apply the following corollary of Lemma 2.5 to a certain submatrix of the coefficient matrix of an essential cover. This allows us to find a point that is far from certain hyperplanes in the cover.

Corollary 2.6. *Let $V \in \mathbb{R}^{\ell \times m}$ be an $\ell \times m$ matrix with row vectors $v_1, \dots, v_\ell \in \mathbb{R}^m$ and column vectors $v_{*1}, \dots, v_{*m} \in \mathbb{R}^\ell$. Suppose that $||v_i||_2 = 1$ for all $i \in [\ell]$, and let $\theta \geq 0$ be a real number satisfying $\theta||v_{*j}||_1 \leq 1/3$ for all $j \in [m]$. Then for any vector $\mu = (\mu_1, \dots, \mu_\ell) \in \mathbb{R}^\ell$, there exists $y \in \mathbb{R}^m$ such that $||y||_\infty \leq 1/3$ and*

$$\big|\langle v_i, y\rangle - \mu_i\big| \geq \theta \quad \text{for all } i \in [\ell].$$

Proof. Apply Lemma 2.5 to the symmetric $\ell \times \ell$ matrix $M := VV^T$ and the vector $\mu \in \mathbb{R}^\ell$. Note that we have $M_{ii} = 1$ for all $i \in [\ell]$ since $||v_i||_2 = 1$ for all $i \in [\ell]$. Hence there exists a vector $\varepsilon \in \{\pm 1\}^\ell$ such that

$$\big|(M(\theta\varepsilon))_i - \mu_i\big| \geq \theta \quad \text{for all } i \in [\ell].$$

Setting $y := \theta V^T \varepsilon \in \mathbb{R}^m$, for every $i \in [\ell]$, we obtain

$$\big|\langle v_i, y\rangle - \mu_i\big| = \big|(Vy)_i - \mu_i\big| = \Big|(V(\theta V^T\varepsilon))_i - \mu_i\Big| = \Big|(VV^T(\theta\varepsilon))_i - \mu_i\Big| = |(M(\theta\varepsilon))_i - \mu_i| \geq \theta.$$

Furthermore, we have

$$||y||_\infty = ||\theta V^T\varepsilon||_\infty = \max_{j\in[m]} \left|\theta \sum_{i=1}^{\ell} v_{ij}\varepsilon_i\right| \leq \max_{j\in[m]} \theta \sum_{i=1}^{\ell} |v_{ij}| = \max_{j\in[m]} \theta||v_{*j}||_1 \leq \frac{1}{3},$$

as desired. □

Finally, we will use the following well-known concentration inequality.

Lemma 2.7 (Hoeffding's inequality [7]). *Let $a_1, \dots, a_\ell, b_1, \dots, b_\ell \in \mathbb{R}$ and let $z_1, \dots, z_\ell$ be independent real random variables such that for all $j \in [\ell]$, we always have $a_j \leq z_j \leq b_j$. Let $z = \sum_{i=1}^{\ell} z_i$, then for every $t > 0$, we have*

$$\mathbb{P}\Big[\big|z - \mathbb{E}[z]\big| \geq t\Big] \leq 2 \cdot \exp\left(-\frac{2t^2}{\sum_{j=1}^{\ell}(b_j - a_j)^2}\right).$$

3. Outline

We employ a strategy similar to the one in [12]. Given an essential cover of the hypercube $\{\pm 1\}^n$, we consider its coefficient matrix, recording the coefficients of the linear equations corresponding to the hyperplanes in the cover. More formally, if the hyperplanes are given by equations of the form $\langle v_i, x\rangle = \mu_i$ for $i = 1, \dots, k$, then the coefficient matrix of the essential cover is the $k \times n$ matrix with rows $v_1, \dots, v_k$.

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

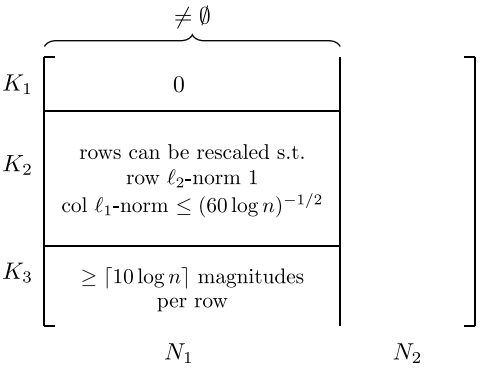


Figure 1. The decomposition of the coefficient matrix V of the essential cover.

We first show that for any essential cover consisting of only a small number of hyperplanes, there is a decomposition of the coefficient matrix of a certain structured form. Using this decomposition, we then obtain a contradiction by probabilistically finding a vertex in $\{\pm 1\}^n$ that is not covered by any of the hyperplanes in the cover. Our decomposition of the coefficient matrix is given by the following proposition.

Proposition 3.1 (Matrix decomposition). *For $n \geq 2$, let $h_1, \ldots, h_k$ be hyperplanes in $\mathbb{R}^n$ forming an essential cover of the hypercube $\{\pm 1\}^n$. For each $i \in [k]$, let $h_i = \{x \in \mathbb{R}^n \mid \langle v_i, x \rangle = \mu_i\}$, and let $V \in \mathbb{R}^{k \times n}$ be the matrix with rows $v_1, \ldots, v_k$. If $k \leq 10^{-2} \cdot n^{2/3}/(\log n)^{2/3}$, then there exists a partition $[k] = K_1 \sqcup K_2 \sqcup K_3$ of the row indices of V and a partition $[n] = N_1 \sqcup N_2$ of the column indices of V with $N_1 \neq \emptyset$, such that the following conditions hold:*

1. *Every entry in the submatrix $V[K_1 \times N_1]$ is zero.*
2. *There exists a $|K_2| \times |N_1|$ matrix V' that can be obtained from $V[K_2 \times N_1]$ by rescaling the rows in some way (by some non-zero real numbers), such that every row in V' has ℓ_2-norm equal to 1 and every column in V' has ℓ_1-norm at most $(60 \log n)^{-1/2}$.*
3. *In the submatrix $V[K_3 \times N_1]$, every row has at least $\lceil 10 \log n \rceil$ magnitudes.*

Figure 1 illustrates the decomposition in Proposition 3.1. We remark that in Proposition 3.1, some of the subsets $K_1, K_2, K_3 \subseteq [k]$ may be empty, and similarly, $N_2 \subseteq [n]$ may be empty.

Given an essential cover whose coefficient matrix can be decomposed as in Proposition 3.1, we show that there must exist a vertex in $\{\pm 1\}^n$ that is not covered by any of the hyperplanes (yielding a contradiction to condition (E1)). This is stated in the following proposition.

Proposition 3.2. *For $n \geq 2$, let $h_1, \ldots, h_k$ be hyperplanes in $\mathbb{R}^n$ forming an essential cover of the hypercube $\{\pm 1\}^n$. For each $i \in [k]$, let $h_i = \{x \in \mathbb{R}^n \mid \langle v_i, x \rangle = \mu_i\}$, and let $V \in \mathbb{R}^{k \times n}$ be the matrix with rows $v_1, \ldots, v_k$. Suppose that the matrix V has a decomposition as in Proposition 3.1 and that $k \leq n$. Then there exists $w \in \{\pm 1\}^n$ such that $\langle v_i, w \rangle \neq \mu_i$ for all $i \in [k]$.*

In other words, this proposition states that actually the coefficient matrix of an essential cover cannot have a decomposition as in Proposition 3.1. This means that we cannot have $k \leq 10^{-2} \cdot n^{2/3}/(\log n)^{2/3}$.

The proof of Proposition 3.2 employs a similar strategy to [3, 12]: we find a certain probability distribution on $\{\pm 1\}^n$ such that with positive probability a random vertex $w \in \{\pm 1\}^n$ with this distribution avoids all the hyperplanes. To do so, we first fix the coordinates of w with indices in N_2 deterministically using the minimality property of an essential cover (property (E3) in Definition 1.1) to always avoid all the hyperplanes h_i with $i \in K_1$. Then we construct probability distributions for the coordinates of w with indices in N_1 such that with positive probability w

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

avoids all the hyperplanes h_i with $i \in K_2 \cup K_3$. The probability distribution is constructed by applying Corollary 2.6 (a corollary of Bang's lemma) to the matrix V' in condition (2) in Proposition 3.1. This way, we find a point $y \in [-1/3, 1/3]^{N_1}$ that is "far away" from all hyperplanes h_i with $i \in K_2$. We then randomly round y to a vertex in $\{\pm 1\}^{N_1}$ of the hypercube to define coordinates of w with indices in N_1. We can show that for any $i \in K_2$, the vertex w is unlikely to be contained in the hyperplane h_i because y is "far away" from h_i. To bound the probability that w is contained in a hyperplane h_i with $i \in K_3$, we use that by condition (3) in Proposition 3.1 every row of the submatrix $V[K_3 \times N_1]$ has many magnitudes. So for each $i \in K_3$, Lemma 2.3 implies an upper bound for the probability that w is contained in h_i.

Despite the similarity in the overall strategy between our argument and the arguments in [3, 12], we highlight the fact that we have a simpler matrix decomposition in Proposition 3.1 and a much shorter proof for Proposition 3.2, while obtaining a better bound. The proof of Proposition 1.2 follows directly from combining Propositions 3.1 and 3.2.

Proof of Theorem 1.2. Suppose for contradiction that the hyperplanes $h_1, \dots, h_k$ form an essential cover of the hypercube $\{\pm 1\}^n$ of size $k < 10^{-2} \cdot n^{2/3}/(\log n)^{2/3}$. For each $i \in [k]$, let $h_i = \{x \in \mathbb{R}^n \mid \langle v_i, x\rangle = \mu_i\}$, and let $V \in \mathbb{R}^{k \times n}$ be the matrix with rows $v_1, \dots, v_k$. Then, we can decompose V as in Proposition 3.1. But now by Proposition 3.2, there must be a vertex $w \in \{\pm 1\}^n$ that is not covered by any of the hyperplanes $h_1, \dots, h_k$. This contradicts condition (E1) in the definition of an essential cover in Definition 1.1. □

The rest of the paper is structured as follows. In section 4, we show that for an essential cover whose coefficient matrix has a decomposition as in Proposition 3.1, there must be a vertex $w \in \{\pm 1\}^n$ that is not covered by any of the hyperplanes; that is, we prove Proposition 3.2. Then in section 5, we prove Proposition 3.1.

4. Finding the uncovered vertex

In this section, we prove Proposition 3.2; that is, we show that given an essential cover of the hypercube $\{\pm 1\}^n$ whose the coefficient matrix $V \in \mathbb{R}^{k \times n}$ can be decomposed as in Proposition 3.1, we can find a vertex $w \in \{\pm 1\}^n$ that is not covered by any of the hyperplanes. For partitions $[k] = K_1 \sqcup K_2 \sqcup K_3$ and $[n] = N_1 \sqcup N_2$ as in Proposition 3.1, we construct $w \in \{\pm 1\}^n$ by first fixing the entries w_j for $j \in N_2$ to avoid all the hyperplanes h_i with $i \in K_1$ deterministically and then probabilistically choosing the entries w_j for $j \in N_1$, avoiding all the hyperplanes h_i with $i \in K_2 \cup K_3$ with positive probability.

Proof of Proposition 3.2. Let $[k] = K_1 \sqcup K_2 \sqcup K_3$ and $[n] = N_1 \sqcup N_2$ be partitions satisfying the conditions in Proposition 3.1. Note that these conditions are not affected by rescaling the hyperplane equations $\langle v_i, x\rangle = \mu_i$ describing the hyperplanes h_i for $i \in K_2$. We may therefore assume that the matrix V' in condition (2) in Proposition 3.1 is simply $V' = V[K_2 \times N_1]$ after rescaling the equations $\langle v_i, x\rangle = \mu_i$ for $i \in K_2$ appropriately (note that this precisely corresponds to rescaling the rows of V with indices in K_2). Specifically, this means that we have $||v_i|_{N_1}||_2 = 1$ for each $i \in K_2$, and furthermore, for each $j \in N_1$, the j-th column v_{*j} of V satisfies $||v_{*j}|_{K_2}||_1 \le (60 \log n)^{-1/2}$.

Next, we claim that $K_1 \neq [k]$. Indeed, condition (E2) in Definition 1.1 means that the matrix V cannot have any column with all entries being zero. On the other hand, by condition (1) in Proposition 3.1 the submatrix $V[K_1 \times N_1]$ of V is the zero matrix. Therefore, if we had $K_1 = [k]$, then for every $j \in N_1$ the j-th column of V would be all-zero. This would be a contradiction, given that $N_1 \neq \emptyset$. Thus, we indeed have $K_1 \neq [k]$.

Now, we can fix the entries w_j for $j \in N_2$ of our desired vertex $w \in \{\pm 1\}^n$ using the following lemma.

Lemma 4.1. *There exists $\tilde{w} \in \{\pm 1\}^{N_2}$ such that for every $w \in \{\pm 1\}^n$ with $w|_{N_2} = \tilde{w}$, we have $\langle v_i, w\rangle \neq \mu_i$ for all $i \in K_1$.*

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

Proof. By condition (E3) in Definition 1.1 applied to some element of $[k] \setminus K_1$, there exists a vertex $x \in \{\pm 1\}^n$ such that $\langle v_i, x\rangle \neq \mu_i$ for all $i \in K_1$. We set $\tilde{w} := x|_{N_2}$. By condition (1) in Proposition 3.1, for every $i \in K_1$, the restriction $v_i|_{N_1}$ is the all-zero vector. Therefore, for every $w \in \{\pm 1\}^n$ with $w|_{N_2} = \tilde{w}$, we have

$$\langle v_i, w\rangle = \langle v_i|_{N_2}, \tilde{w}\rangle = \langle v_i|_{N_2}, x|_{N_2}\rangle = \langle v_i, x\rangle \neq \mu_i$$

for all $i \in K_1$, as desired. □

Let $\tilde{w} \in \{\pm 1\}^{N_2}$ be as in Lemma 4.1, and fix $w_j = \tilde{w}_j$ for all $j \in N_2$. Next, we define a probability distribution for the entries w_j for $j \in N_1$. To do so, we apply Corollary 2.6 to the matrix, $V[K_2 \times N_1]$ with $\theta = (6 \log n)^{1/2}$ and $\mu' \in \mathbb{R}^{K_2}$ defined by $\mu'_i = \mu_i - \langle v_i|_{N_2}, \tilde{w}\rangle$ for all $j \in K_2$. We have $\theta ||v_{*j}|_{K_2}||_1 \leq (6 \log n)^{1/2} \cdot (60 \log n)^{-1/2} \leq 1/3$ for all $j \in N_1$, as well as $||v_i|_{N_1}||_2 = 1$ for all $i \in K_2$, so the conditions of Corollary 2.6 are satisfied for the matrix $V[K_2 \times N_1]$. So we obtain $y \in \mathbb{R}^{N_1}$ such that $||y||_\infty \leq 1/3$ and $\left|\langle v_i|_{N_1}, y\rangle - \mu'_i\right| \geq \theta$ for all $i \in K_2$. Then for every $j \in N_1$, let us take a random variable $w_j \in \{\pm 1\}$ with distribution

$$w_j = \begin{cases} 1 & \text{with probability } (1 + y_j)/2 \\ -1 & \text{with probability } (1 - y_j)/2, \end{cases} \tag{1}$$

such that the random variables w_j are independent for all $j \in N_1$. Note that then we have $\mathbb{E}[w_j] = y_j$ for each $j \in N_1$.

Taking the entries of w to be the random variables w_j for $j \in N_1$ together with the fixed values $w_j = \tilde{w}_j$ for $j \in N_2$ defined above, we have now defined a random vertex $w \in \{\pm 1\}^n$. Our goal is to show that with positive probability $w \in \{\pm 1\}^n$ satisfies $\langle v_i, w\rangle \neq \mu_i$ for all $i \in [k]$. By the choice of $\tilde{w}$ in Lemma 4.1 we always have $\langle v_i, w\rangle \neq \mu_i$ for all $i \in K_1$. Next, we show that for $i \in K_2$ we are unlikely to have $\langle w, v_i\rangle = \mu_i$.

Claim 4.2. *For every $i \in K_2$, we have*

$$\mathbb{P}\Big[\langle v_i, w\rangle = \mu_i\Big] \leq \frac{1}{2n}.$$

Proof. Fix $i \in K_2$. By our definition of $\mu'_i = \mu_i - \langle v_i|_{N_2}, \tilde{w}\rangle = \mu_i - \langle v_i|_{N_2}, w|_{N_2}\rangle$, we have $\langle v_i, w\rangle = \mu_i$ if and only if $\langle v_i|_{N_1}, w|_{N_1}\rangle = \mu'_i$. So it suffices to show that

$$\mathbb{P}\Big[\langle v_i|_{N_1}, w|_{N_1}\rangle = \mu'_i\Big] \leq \frac{1}{2n}.$$

Since $\mathbb{E}[w_j] = y_j$ for all $j \in N_1$, by linearity of expectation, we have $\mathbb{E}[\langle v_i|_{N_1}, w|_{N_1}\rangle] = \langle v_i|_{N_1}, y\rangle$. In addition, recall that we have $\left|\langle v_i|_{N_1}, y\rangle - \mu'_i\right| \geq \theta$, so we obtain

$$\begin{aligned}\mathbb{P}\Big[\langle v_i|_{N_1}, w|_{N_1}\rangle = \mu'_i\Big] &\leq \mathbb{P}\Big[\left|\langle v_i|_{N_1}, w|_{N_1}\rangle - \langle v_i|_{N_1}, y\rangle\right| \geq \theta\Big] \\ &= \mathbb{P}\Big[\left|\langle v_i|_{N_1}, w|_{N_1}\rangle - \mathbb{E}[\langle v_i|_{N_1}, w|_{N_1}\rangle]\right| \geq \theta\Big].\end{aligned}$$

Now, we can apply Hoeffding's inequality (Lemma 2.7) to the random variables $v_{ij}w_j$ for $j \in N_1$, whose sum is $\sum_{j \in N_1} v_{ij}w_j = \langle v_i|_{N_1}, w|_{N_1}\rangle$. Note that the random variables $v_{ij}w_j$ are independent and bounded by $-|v_{ij}| \leq v_{ij}w_j \leq |v_{ij}|$ for each $j \in N_1$. Since we have $||v_i|_{N_1}||_2 = 1$, we obtain

$$\begin{aligned}\mathbb{P}\Big[\left|\langle v_i|_{N_1}, w|_{N_1}\rangle - \mathbb{E}[\langle v_i|_{N_1}, w|_{N_1}\rangle]\right| \geq \theta\Big] &\leq 2 \cdot \exp\left(-\frac{2\theta^2}{\sum_{j \in N_1}(2v_{ij})^2}\right) \\ &= 2 \cdot \exp\left(-\frac{2\theta^2}{4||v_i|_{N_1}||_2^2}\right) \\ &= 2 \cdot \exp\left(-\frac{\theta^2}{2}\right).\end{aligned}$$

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

Substituting in our choice of $\theta = (6 \log n)^{1/2}$, we can deduce (recalling $n \geq 2$)

$$\mathbb{P}\Big[\langle v_i|_{N_1}, w|_{N_1}\rangle = \mu_i'\Big] \leq 2\exp(-3\log n) = \frac{2}{n^3} \leq \frac{1}{2n}.$$

□

On the other hand, by condition (3) in Proposition 3.1, for every $i \in K_3$, the vector $v_i|_{N_1}$ has at least $\lceil 10 \log n \rceil$ magnitudes. Thus, by Lemma 2.3 applied to $p \in [1/3, 2/3]^{N_1}$ given by $p_j = (1 + y_j)/2$ for all $j \in N_1$, we have

$$\mathbb{P}\Big[\langle v_i, w\rangle = \mu_i\Big] = \mathbb{P}\Big[\langle v_i|_{N_1}, w|_{N_1}\rangle = \mu_i - \langle v_i|_{N_2}, \tilde{w}\rangle\Big] \leq \left(\frac{2}{3}\right)^{\lceil 10\log n/2\rceil} \leq \left(\frac{2}{3}\right)^{5\log n} \leq \frac{1}{n^2} \leq \frac{1}{2n}$$

for every $i \in K_3$. Together with Claim 4.2, we can take a union bound over all $i \in K_2 \cup K_3$ and obtain

$$\mathbb{P}\Big[\langle v_i, w\rangle = \mu_i \text{ for some } i \in K_2 \cup K_3\Big] \leq \sum_{i \in K_2 \cup K_3} \mathbb{P}\Big[\langle v_i, w\rangle = \mu_i\Big] \leq k \cdot \frac{1}{2n} \leq n \cdot \frac{1}{2n} = \frac{1}{2},$$

recalling that $k \leq n$.

Thus, with probability at least $1/2$, our random vertex $w \in \{\pm 1\}^n$ satisfies $\langle v_i, w\rangle \neq \mu_i$ for all $i \in K_2 \cup K_3$. Recalling that we always have $\langle v_i, w\rangle \neq \mu_i$ for all $i \in K_1$, this means that with probability at least $1/2$, we have $\langle v_i, w\rangle \neq \mu_i$ for all $i \in [k]$. Hence, there exists a vertex in $w \in \{\pm 1\}^n$ such that $\langle v_i, w\rangle \neq \mu_i$ for all $i \in [k]$. □

5. Decomposing the matrix

Finally, we prove Proposition 3.1.

Proof of Proposition 3.1. We decompose the matrix V using the following algorithm that starts with all the column indices in N_1. At any iteration of the algorithm, we have a partition $[n] = N_1 \sqcup N_2$ of the column indices, and furthermore, the row indices $i \in [k]$ are partitioned into three types: being in K_1, active, and inactive (defined later). Throughout the algorithm, we move certain columns from N_1 to N_2. When the algorithm terminates, the desired partitions of the row and column indices of the matrix will be formed by taking the sets K_1, N_1, and N_2 at the end of the algorithm, as well as taking K_2 to be the active row indices and K_3 to be the inactive row indices at the end of the algorithm.

1. Initialise $N_1 = [n]$, $N_2 = \emptyset$ and $K_1 = \emptyset$.
2. For each $i \in [k] \setminus K_1$ such that $v_i|_{N_1}$ is the all-zero vector, we move i into K_1.
3. For each $i \in [k] \setminus K_1$, we say that i is *inactive* if the vector $v_i|_{N_1}$ has at least $\lceil 10 \log n \rceil$ magnitudes; otherwise, we say that i is *active*.
4. For each active $i \in [k] \setminus K_1$ and each $j \in N_1$, we assign a weight w_{ij} to the entry v_{ij} of V as follows: if $v_{ij} = 0$, define $w_{ij} = 0$. If $v_{ij} \neq 0$, let m be the magnitude of v_{ij}, and let $N_m(i)$ be the number of entries of $v_i|_{N_1}$ with magnitude m, and define $w_{ij} = 10/\sqrt{N_m(i)}$.
5. For every $j \in N_1$, calculate the total weight $\sum_i w_{ij}$ in column j, where the sum is taken over all active $i \in [k] \setminus K_1$. If there exists some $j \in N_1$ where this total weight is larger than $(60 \log n)^{-1/2}$, then move this index j from N_1 to N_2 and remove all weights in column j (if there exist multiple such j, choose one arbitrarily), and go to back to Step 2. If there is no such index $j \in N_1$, go to Step 6.

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

6. Terminate and output the current sets N_1, N_2, K_1, output the current set of active row indices as K_2, and output the current set of inactive row indices as K_3.

It is clear that this algorithm must eventually terminate since in every iteration, a column index is moved from N_1 to N_2 (so the set N_1 is getting strictly smaller with every iteration). It is also clear that the algorithm produces a partition $[k] = K_1 \sqcup K_2 \sqcup K_3$ of the row indices of V and a partition $[n] = N_1 \sqcup N_2$ of the column indices. Furthermore, it is not hard to see that these partitions satisfy conditions (1) and (3) in Proposition 3.1. Indeed, condition (1) is satisfied by the rule of moving row indices into K_1 in Step 2 of the algorithm, and condition (3) is satisfied by the definition of inactive rows in Step 3.

To show condition (2) in Proposition 3.1, consider the $|K_2| \times |N_1|$ matrix V' obtained from $V[K_2 \times N_1]$ by normalising every row to have ℓ_2-norm equal to 1 (note that by the condition in Step 2, all the rows of $V[K_2 \times N_1]$ are non-zero). Let us write $v'_i = (1/||v_i|_{N_1}||_2)v_i|_{N_1}$ for $i \in K_2$ for the rows of V'. Now, for every entry $v_{ij} \neq 0$ of $V[K_2 \times N_1]$ with some magnitude m, we have $10^m \leq |v_{ij}| < 10^{m+1}$, and for each $i \in K_2$, there are $N_m(i)$ such entries in $v_i|_{N_1}$. Thus, $||v_i|_{N_1}||_2 \geq \sqrt{N_m(i)} \cdot 10^m$ for every magnitude m appearing among the entries of $v_i|_{N_1}$, and we can conclude that

$$|v'_{ij}| = \frac{|v_{ij}|}{||v_i|_{N_1}||_2} < \frac{10^{m+1}}{\sqrt{N_m(i)} \cdot 10^m} = \frac{10}{\sqrt{N_m(i)}} = w_{ij}$$

for every entry $v_{ij} \neq 0$ of $V[K_2 \times N_1]$ with magnitude m. This shows that $|v'_{ij}| \leq w_{ij}$ for all $i \in K_2$ and $j \in N_1$ (note that in the case $v_{ij} = 0$, we trivially have $|v'_{ij}| = 0 = w_{ij}$).

Now, to check condition (2), observe that for every $j \in N_1$, we have $\sum_{i \in K_2} w_{ij} \leq (60 \log n)^{-1/2}$ since otherwise the algorithm would not have terminated. Consequently, for every $j \in N_1$, we have $\sum_{i \in K_2} |v'_{ij}| \leq \sum_{i \in K_2} w_{ij} \leq (60 \log n)^{-1/2}$, meaning that the ℓ_1-norm of every column in V' is at most $(60 \log n)^{-1/2}$.

It remains to show $N_1 \neq \emptyset$. Recall that at the start, we have $N_1 = [n]$, and in every iteration of the algorithm, we move one index from N_1 into N_2. So we need to show that the number of column indices moved into N_2 throughout the algorithm is less than n. At every step, we consider the total weight in the matrix. Whenever we move a column index from N_1 into N_2, this total weight decreases by at least $(60 \log n)^{-1/2}$ (by the condition on the column weight in Step 5 of the algorithm).

However, note after moving a column from N_1 to N_2, when reassigning the weights in Step 3 in the next iteration, some of the weights in the remaining columns with indices in N_1 may increase (indeed, the numbers $N_m(i)$ may decrease by removing a column from N_1). Furthermore, new row indices may become active as the set N_1 gets smaller throughout the algorithm (but note that once a row index becomes active, it cannot become inactive anymore afterwards).

For every row index $i \in [k]$ that becomes active at some point during the algorithm, and every magnitude occurring among the entries of $v_i|_{N_1}$ at the moment when the index i first becomes active, let $N^*_m(i)$ denote the number of entries of $v_i|_{N_1}$ with magnitude m when the row index i first becomes active. At that iteration of the algorithm, each of these $N^*_m(i)$ entries is assigned weight $10/\sqrt{N^*_m(i)}$. When a column index of some entry of $v_i|_{N_1}$ with magnitude m gets moved into N_2 later in the algorithm, the quantity $N_m(i)$ decreases, and so the weights of the remaining entries of $v_i|_{N_1}$ with magnitude m increase. More precisely, when this happens for the t-th time (for $t \in \{1, \ldots, N^*_m(i) - 1\}$), the weight of the $N^*_m(i) - t$ remaining entries of $v_i|_{N_1}$ with magnitude m is increased from $10/\sqrt{N^*_m(i) - t + 1}$ to $10/\sqrt{N^*_m(i) - t}$. Thus, the total weight that is added to the entries with magnitude m in row i throughout the algorithm (both when row i first becomes active and as the weights get re-assigned) is at most

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

$$N_m^*(i)\cdot\frac{10}{\sqrt{N_m^*(i)}}+\sum_{t=1}^{N_m^*(i)-1}\left(N_m^*(i)-t\right)\left(\frac{10}{\sqrt{N_m^*(i)-t}}-\frac{10}{\sqrt{N_m^*(i)-t+1}}\right)$$
$$=10\sqrt{N_m^*(i)}+10\sum_{t=1}^{N_m^*(i)-1}t\left(\frac{1}{\sqrt{t}}-\frac{1}{\sqrt{t+1}}\right)=10\sqrt{N_m^*(i)}+10\sum_{t=1}^{N_m^*(i)-1}t\cdot\frac{\sqrt{t+1}-\sqrt{t}}{\sqrt{t(t+1)}}$$
$$\le 10\sqrt{N_m^*(i)}+10\sum_{t=1}^{N_m^*(i)-1}\left(\sqrt{t+1}-\sqrt{t}\right)\le 10\sqrt{N_m^*(i)}+10\sqrt{N_m^*(i)}=20\sqrt{N_m^*(i)}$$

for any row index i becoming active at some point during the algorithm and any m appearing as magnitude among the entries of $v_i|_{N_1}$ at the moment when the index i first becomes active.

For notational convenience, let $S=\lceil 10\log n\rceil-1$. Recall that when the row index i becomes active, there are strictly less than $\lceil 10\log n\rceil$ (i.e., at most S) magnitudes appearing among the entries of $v_i|_{N_1}$. For these at most S different magnitudes m, the numbers $N_m^*(i)$ defined above (counting the entries of $v_i|_{N_1}$ of magnitude m when the row index i first becomes active) satisfy

$$\sum_m N_m^*(i)=|\operatorname{supp}(v_i|_{N_1})|\le|\operatorname{supp}(v_i)|\le 2k,$$

where the last inequality holds by Lemma 2.4. Hence

$$\sum_m 20\sqrt{N_m(i)}=20\sum_m\sqrt{N_m(i)}\le 20\cdot\sqrt{S}\cdot\sqrt{\sum_m N_m^*(i)}\le 20\cdot\sqrt{S}\cdot\sqrt{2k}\le 30\sqrt{kS}$$

by the Cauchy–Schwarz inequality, so for every row, the total weight added to the entries throughout the algorithm is at most $30\sqrt{kS}$. Thus, the total weight increase for the entire $k\times n$ matrix V throughout the algorithm is at most

$$k\cdot 30\sqrt{kS}=30k^{3/2}\sqrt{S}\le 30k^{3/2}\cdot(10\log n)^{1/2}.$$

Recall that whenever we move a column index from N_1 into N_2, the weight in the matrix V decreases by at least $(60\log n)^{-1/2}$. Therefore, at most

$$\frac{30k^{3/2}\cdot(10\log n)^{1/2}}{(60\log n)^{-1/2}}\le 750k^{3/2}\cdot\log n$$

column indices can be moved from N_1 into N_2. By our assumption $k\le 10^{-2}\cdot n^{2/3}/(\log n)^{2/3}$, this number of column indices moved from N_1 into N_2 is at most

$$750k^{3/2}\cdot\log n\le 750\cdot 10^{-3}\cdot\frac{n}{\log n}\cdot\log n<n.$$

Thus, we have $N_1\neq\emptyset$ at the end of the algorithm, as desired. □

References

[1] Alon, N. (1999) Combinatorial Nullstellensatz. *Comb. Probab. Comp.* **8**(1-2) 7–29. DOI: 10.1017/S0963548398003411.

[2] Alon, N. and Füredi, Z. (1993) Covering the cube by affine hyperplanes. *European J. Combin.* **14**(2) 79–83. DOI: 10.1006/eujc.1993.1011.

[3] Araujo, I., Balogh, J. and Mattos, L. (2022) New lower bounds for essential covers of the cube, arXiv:2209.00140.

[4] Ball, K. (1991) The plank problem for symmetric bodies. *Invent. Math.* **104**(3) 535–543. DOI: 10.1007/BF01245089.

[5] Bang, T. (1951) A solution of the "plank problem. *Proc. Amer. Math. Soc.* **2**(6) 990–993. DOI: 10.2307/2031721.

[6] Clifton, A. and Huang, H. (2020) On almost k-covers of hypercubes. *Combinatorica* **40**(4) 511–526. DOI: 10.1007/s00493-019-4221-y.

[7] Hoeffding, W. (1963) Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**(301) 13–30. DOI: 10.2307/2282952.

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press

[8] Ivanisvili, P., Klein, O. and Vershynin, R. (2023) Covering the hypercube, the uncertainty principle, and an interpolation formula, arXiv:2310.13277.
[9] Klein, O. (2023) Slicing all Edges of an n-cube Requires $n^{2/3}$ Hyperplanes, pp. 1931–1936. In: 2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS), DOI: 10.1109/FOCS57990.2023.00117.
[10] Linial, N. and Radhakrishnan, J. (2005) Essential covers of the cube by hyperplanes. *J. Comb. Theory Ser. A* **109**(2) 331–338. DOI: 10.1016/j.jcta.2004.07.012.
[11] Saxton, D. (2013) Essential positive covers of the cube. *J. Comb. Theory Ser. A* **120**(5) 971–975. DOI: 10.1016/j.jcta.2013.02.002.
[12] Yehuda, G. and Yehudayoff, A. (2024) A lower bound for essential covers of the cube. *Combinatorica* **44**(4) 801–815. DOI: 10.1007/s00493-024-00093-4.
[13] Yehuda, G. and Yehudayoff, A. (2021) Slicing the hypercube is not easy, arXiv:2102.05536.

Cite this article: Sauermann L and Xu Z (2025). Essential covers of the hypercube require many hyperplanes. *Combinatorics, Probability and Computing* **34**, 326–337. https://doi.org/10.1017/S0963548324000257

https://doi.org/10.1017/S0963548324000257 Published online by Cambridge University Press